

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/519,069  
Source: pg 10  
Date Processed by STIC: 11/3/05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 11/03/2005

PATENT APPLICATION: US/10/519,069

TIME: 08:23:18

Input Set : A:\Seqlist.txt

Output Set: N:\CRF4\11032005\J519069.raw

4 <110> APPLICANT: Fritsch Olivier  
 5 Hohn Barbara  
 6 Lucht Jan Martin  
 8 <120> TITLE OF INVENTION: GENE FOR INCREASED SOMATIC RECOMBINATION  
 11 <130> FILE REFERENCE: 1-32546B/FMI  
 13 <140> CURRENT APPLICATION NUMBER: 10/519,069  
 C--> 14 <141> CURRENT FILING DATE: 2004-12-22  
 16 <150> PRIOR APPLICATION NUMBER: PCT/EP03/006757  
 17 <151> PRIOR FILING DATE: 2003-06-26  
 19 <150> PRIOR APPLICATION NUMBER: GB 0214896  
 20 <151> PRIOR FILING DATE: 2002-06-27  
 22 <160> NUMBER OF SEQ ID NOS: 30  
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 4524  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Arabidopsis thaliana  
 31 <220> FEATURE:  
 32 <221> NAME/KEY: CDS  
 33 <222> LOCATION: (1)...(4524)  
 34 <223> OTHER INFORMATION: Atino80 sequence  
 36 <400> SEQUENCE: 1  
 37 atg gat cct tca aga cga cca ccg aag gac tct cct tac gcg aat cta 48  
 38 Met Asp Pro Ser Arg Arg Pro Pro Lys Asp Ser Pro Tyr Ala Asn Leu  
 39 1 5 10 15  
 41 ttc gat ctc gag ccg ttg atg aag ttt aga att ccg aaa cct gaa gat 96  
 42 Phe Asp Leu Glu Pro Leu Met Lys Phe Arg Ile Pro Lys Pro Glu Asp  
 43 20 25 30  
 45 gaa gtt gat tat tat ggg agt agt agc cag gat gaa agt aga agc act 144  
 46 Glu Val Asp Tyr Tyr Gly Ser Ser Ser Gln Asp Glu Ser Arg Ser Thr  
 47 35 40 45  
 49 caa ggt ggg gta gtg gca aac tac agc aat ggg tct aaa tcg aga atg 192  
 50 Gln Gly Gly Val Val Ala Asn Tyr Ser Asn Gly Ser Lys Ser Arg Met  
 51 50 55 60  
 53 aat gcg agc tcc aag aag aga aag cgg tgg aca gaa gct gag gat gca 240  
 54 Asn Ala Ser Ser Lys Lys Arg Lys Arg Trp Thr Glu Ala Glu Asp Ala  
 55 65 70 75 80  
 57 gag gac gat gat gat ctc tac aat caa cat gtt act gag gag cac tac 288  
 58 Glu Asp Asp Asp Asp Leu Tyr Asn Gln His Val Thr Glu Glu His Tyr  
 59 85 90 95  
 61 cga tca atg ctt ggg gag cat gta caa aaa ttc aaa aat agg tcc aag 336  
 62 Arg Ser Met Leu Gly Glu His Val Gln Lys Phe Lys Asn Arg Ser Lys  
 63 100 105 110

## RAW SEQUENCE LISTING

DATE: 11/03/2005

PATENT APPLICATION: US/10/519,069

TIME: 08:23:18

Input Set : A:\Seqlist.txt

Output Set: N:\CRF4\11032005\J519069.raw

```

65 gag act caa ggg aat cct cct cat ctg atg ggt ttt ccg gtg cta aag 384
66 Glu Thr Gln Gly Asn Pro Pro His Leu Met Gly Phe Pro Val Leu Lys
67      115      120      125
69 agc aat gtg ggc agt tac aga ggt agg aaa cca ggg aat gat tac cat 432
70 Ser Asn Val Gly Ser Tyr Arg Gly Arg Lys Pro Gly Asn Asp Tyr His
71      130      135      140
73 ggg agg ttc tat gac atg gac aac tct cca aat ttt gca gct gat gtg 480
74 Gly Arg Phe Tyr Asp Met Asp Asn Ser Pro Asn Phe Ala Ala Asp Val
75 145      150      155      160
77 acc cca cat agg cga gga agc tac cat gat cgt gat att aca ccc aag 528
78 Thr Pro His Arg Arg Gly Ser Tyr His Asp Arg Asp Ile Thr Pro Lys
79      165      170      175
81 ata gca tat gaa cct tcg tat ttg gac att ggt gat ggt gtc atc tac 576
82 Ile Ala Tyr Glu Pro Ser Tyr Leu Asp Ile Gly Asp Gly Val Ile Tyr
83      180      185      190
85 aaa atc ccc cca agt tat gac aag ctg gtg gca tca tta aac tta ccg 624
86 Lys Ile Pro Pro Ser Tyr Asp Lys Leu Val Ala Ser Leu Asn Leu Pro
87      195      200      205
89 agc ttt tca gac att cat gtg gaa gaa ttt tac ttg aaa gga act ctg 672
90 Ser Phe Ser Asp Ile His Val Glu Glu Phe Tyr Leu Lys Gly Thr Leu
91      210      215      220
93 gat ctg aga tca tta gca gaa ctg atg gca agt gat aaa agg tct gga 720
94 Asp Leu Arg Ser Leu Ala Glu Leu Met Ala Ser Asp Lys Arg Ser Gly
95 225      230      235      240
97 gta aga agc cgt aat gga atg ggt gag cct cga cct caa tat gaa tct 768
98 Val Arg Ser Arg Asn Gly Met Gly Glu Pro Arg Pro Gln Tyr Glu Ser
99      245      250      255
101 ctt caa gct aga atg aag gcc ctg tca cct tca aac tcc acc cca aat 816
102 Leu Gln Ala Arg Met Lys Ala Leu Ser Pro Ser Asn Ser Thr Pro Asn
103      260      265      270
105 ttt agc ctc aag gtg tca gaa gct gca atg aat tct gcc att cca gaa 864
106 Phe Ser Leu Lys Val Ser Glu Ala Ala Met Asn Ser Ala Ile Pro Glu
107      275      280      285
109 gga tct gct gga agt act gca cgg aca att ctg tct gag ggt ggt gtt 912
110 Gly Ser Ala Gly Ser Thr Ala Arg Thr Ile Leu Ser Glu Gly Gly Val
111      290      295      300
113 tta cag gtc cat tac gtg aag att ctg gag aag ggg gat aca tac gag 960
114 Leu Gln Val His Tyr Val Lys Ile Leu Glu Lys Gly Asp Thr Tyr Glu
115 305      310      315      320
117 att gtt aaa cga agt cta ccg aag aag ctg aaa gca aag aat gat cct 1008
118 Ile Val Lys Arg Ser Leu Pro Lys Lys Leu Lys Ala Lys Asn Asp Pro
119      325      330      335
121 gca gtc att gag aaa aca gaa agg gat aaa att aga aaa gcc tgg atc 1056
122 Ala Val Ile Glu Lys Thr Glu Arg Asp Lys Ile Arg Lys Ala Trp Ile
123      340      345      350
125 aat att gtc aga aga gat ata gca aaa cac cat aga att ttc act act 1104
126 Asn Ile Val Arg Arg Asp Ile Ala Lys His His Arg Ile Phe Thr Thr
127      355      360      365
129 ttt cat cgt aaa cta tca att gat gcc aag agg ttt gca gat ggt tgc 1152

```

## RAW SEQUENCE LISTING

DATE: 11/03/2005

PATENT APPLICATION: US/10/519,069

TIME: 08:23:18

Input Set : A:\Seqlist.txt

Output Set: N:\CRF4\11032005\J519069.raw

```

130 Phe His Arg Lys Leu Ser Ile Asp Ala Lys Arg Phe Ala Asp Gly Cys
131      370      375      380
133 caa aga gag gtg aga atg aag gtg ggt aga tca tac aaa atc cca aga 1200
134 Gln Arg Glu Val Arg Met Lys Val Gly Arg Ser Tyr Lys Ile Pro Arg
135 385      390      395      400
137 act gca cca att cgc act agg aag ata tcc aga gac atg ctg cta ttc 1248
138 Thr Ala Pro Ile Arg Thr Arg Lys Ile Ser Arg Asp Met Leu Leu Phe
139      405      410      415
141 tgg aag cga tat gac aag cag atg gca gaa gag agg aaa aag caa gaa 1296
142 Trp Lys Arg Tyr Asp Lys Gln Met Ala Glu Glu Arg Lys Lys Gln Glu
143      420      425      430
145 aag gaa gct gca gag gct ttt aaa cgt gaa cag gag cag cga gag tca 1344
146 Lys Glu Ala Ala Glu Ala Phe Lys Arg Glu Gln Glu Gln Arg Glu Ser
147      435      440      445
149 aaa agg cag caa caa agg ctc aat ttc ctt att aaa cag act gag ctt 1392
150 Lys Arg Gln Gln Gln Arg Leu Asn Phe Leu Ile Lys Gln Thr Glu Leu
151      450      455      460
153 tac agt cac ttc atg caa aac aag acc gat tgc aat cct tcc gaa gcc 1440
154 Tyr Ser His Phe Met Gln Asn Lys Thr Asp Ser Asn Pro Ser Glu Ala
155 465      470      475      480
157 tta cca ata ggt gat gaa aat ccg att gac gaa gtg ctc cca gaa act 1488
158 Leu Pro Ile Gly Asp Glu Asn Pro Ile Asp Glu Val Leu Pro Glu Thr
159      485      490      495
161 tca gcg gca gaa cct tct gag gta gag gat cct gaa gag gct gaa ctg 1536
162 Ser Ala Ala Glu Pro Ser Glu Val Glu Asp Pro Glu Glu Ala Glu Leu
163      500      505      510
165 aag gaa aag gtc ttg aga gct gcc caa gat gcg gtg tct aag cag aag 1584
166 Lys Glu Lys Val Leu Arg Ala Ala Gln Asp Ala Val Ser Lys Gln Lys
167      515      520      525
169 caa ata aca gat gca ttt gac act gaa tat atg aag cta cgc caa act 1632
170 Gln Ile Thr Asp Ala Phe Asp Thr Glu Tyr Met Lys Leu Arg Gln Thr
171      530      535      540
173 tct gaa atg gaa ggt cct tta aat gat ata tca gtt tct ggc tgc agc 1680
174 Ser Glu Met Glu Gly Pro Leu Asn Asp Ile Ser Val Ser Gly Ser Ser
175 545      550      555      560
177 aat ata gat ttg cat aac cca tct aca atg cct gtt aca tca aca gtt 1728
178 Asn Ile Asp Leu His Asn Pro Ser Thr Met Pro Val Thr Ser Thr Val
179      565      570      575
181 cag act cca gag tta ttt aaa gga acc ctt aaa gaa tac caa atg aaa 1776
182 Gln Thr Pro Glu Leu Phe Lys Gly Thr Leu Lys Glu Tyr Gln Met Lys
183      580      585      590
185 ggc ctt cag tgg cta gtc aat tgt tat gag cag ggt ttg aat ggc ata 1824
186 Gly Leu Gln Trp Leu Val Asn Cys Tyr Glu Gln Gly Leu Asn Gly Ile
187      595      600      605
189 ctt gct gat gaa atg ggc ttg ggt aag act att caa gct atg gcg ttc 1872
190 Leu Ala Asp Glu Met Gly Leu Gly Lys Thr Ile Gln Ala Met Ala Phe
191      610      615      620
193 ttg gca cat ttg gct gag gaa aag aac att tgg ggt cca ttt ctt gtt 1920
194 Leu Ala His Leu Ala Glu Glu Lys Asn Ile Trp Gly Pro Phe Leu Val

```

## RAW SEQUENCE LISTING

DATE: 11/03/2005

PATENT APPLICATION: US/10/519,069

TIME: 08:23:18

Input Set : A:\Seqlist.txt

Output Set: N:\CRF4\11032005\J519069.raw

195	625				630				635				640				
197	gtt	gcc	cct	gcc	tct	gtt	ctt	aac	aat	tgg	gct	gat	gaa	atc	agt	cgt	1968
198	Val	Ala	Pro	Ala	Ser	Val	Leu	Asn	Asn	Trp	Ala	Asp	Glu	Ile	Ser	Arg	
199					645					650					655		
201	ttc	tgt	cct	gac	ttg	aaa	act	ctt	cca	tat	tgg	gga	gga	tta	caa	gaa	2016
202	Phe	Cys	Pro	Asp	Leu	Lys	Thr	Leu	Pro	Tyr	Trp	Gly	Gly	Leu	Gln	Glu	
203				660					665					670			
205	cga	aca	att	tta	aga	aag	aat	atc	aat	ccc	aag	cgt	atg	tac	cga	agg	2064
206	Arg	Thr	Ile	Leu	Arg	Lys	Asn	Ile	Asn	Pro	Lys	Arg	Met	Tyr	Arg	Arg	
207			675					680					685				
209	gat	gct	ggc	ttt	cat	att	ttg	att	act	agc	tat	cag	cta	tta	gtc	act	2112
210	Asp	Ala	Gly	Phe	His	Ile	Leu	Ile	Thr	Ser	Tyr	Gln	Leu	Leu	Val	Thr	
211		690					695					700					
213	gat	gaa	aag	tat	ttt	cgc	cgg	gtg	aag	tgg	caa	tat	atg	gtg	cta	gat	2160
214	Asp	Glu	Lys	Tyr	Phe	Arg	Arg	Val	Lys	Trp	Gln	Tyr	Met	Val	Leu	Asp	
215	705					710				715				720			
217	gag	gcc	caa	gca	atc	aag	agt	tcc	tcc	agt	ata	aga	tgg	aaa	acc	ctt	2208
218	Glu	Ala	Gln	Ala	Ile	Lys	Ser	Ser	Ser	Ser	Ile	Arg	Trp	Lys	Thr	Leu	
219				725					730				735				
221	ctt	agt	ttt	aac	tgt	cgg	aac	cga	ttg	ctt	ctg	act	ggg	act	cca	att	2256
222	Leu	Ser	Phe	Asn	Cys	Arg	Asn	Arg	Leu	Leu	Leu	Thr	Gly	Thr	Pro	Ile	
223			740					745					750				
225	cag	aac	aac	atg	gca	gag	tta	tgg	gcc	ctg	ctg	cat	ttc	atc	atg	cca	2304
226	Gln	Asn	Asn	Met	Ala	Glu	Leu	Trp	Ala	Leu	Leu	His	Phe	Ile	Met	Pro	
227		755					760					765					
229	atg	ttg	ttt	gac	aac	cat	gat	caa	ttt	aat	gaa	tgg	ttc	tca	aaa	gga	2352
230	Met	Leu	Phe	Asp	Asn	His	Asp	Gln	Phe	Asn	Glu	Trp	Phe	Ser	Lys	Gly	
231		770				775				780							
233	att	gag	aat	cat	gct	gaa	cac	gga	ggc	act	tta	aat	gag	cac	cag	ctt	2400
234	Ile	Glu	Asn	His	Ala	Glu	His	Gly	Gly	Thr	Leu	Asn	Glu	His	Gln	Leu	
235	785				790				795				800				
237	aac	aga	ctg	cat	gcg	atc	ttg	aaa	ccg	ttc	atg	ctt	cga	cgg	gta	aaa	2448
238	Asn	Arg	Leu	His	Ala	Ile	Leu	Lys	Pro	Phe	Met	Leu	Arg	Arg	Val	Lys	
239				805				810					815				
241	aag	gat	gtg	gtt	tct	gag	cta	act	aca	aag	acg	gaa	gtt	aca	gta	cac	2496
242	Lys	Asp	Val	Val	Ser	Glu	Leu	Thr	Thr	Lys	Thr	Glu	Val	Thr	Val	His	
243			820					825					830				
245	tgc	aag	ctc	agt	tct	cga	caa	caa	gct	ttt	tat	cag	gct	att	aag	aac	2544
246	Cys	Lys	Leu	Ser	Ser	Arg	Gln	Gln	Ala	Phe	Tyr	Gln	Ala	Ile	Lys	Asn	
247		835					840					845					
249	aaa	att	tct	ctg	gct	gag	ttg	ttt	gat	agc	aac	cgc	gga	caa	ttt	act	2592
250	Lys	Ile	Ser	Leu	Ala	Glu	Leu	Phe	Asp	Ser	Asn	Arg	Gly	Gln	Phe	Thr	
251		850				855						860					
253	gat	aag	aaa	gta	ttg	aat	tta	atg	aat	att	gtc	att	caa	cta	agg	aag	2640
254	Asp	Lys	Lys	Val	Leu	Asn	Leu	Met	Asn	Ile	Val	Ile	Gln	Leu	Arg	Lys	
255	865				870				875				880				
257	gtt	tgc	aac	cat	cca	gag	ttg	ttc	gaa	agg	aat	gaa	ggg	agc	tcg	tat	2688
258	Val	Cys	Asn	His	Pro	Glu	Leu	Phe	Glu	Arg	Asn	Glu	Gly	Ser	Ser	Tyr	
259				885				890					895				

## RAW SEQUENCE LISTING

DATE: 11/03/2005

PATENT APPLICATION: US/10/519,069

TIME: 08:23:18

Input Set : A:\Seqlist.txt

Output Set: N:\CRF4\11032005\J519069.raw

```

261 ctc tac ttt gga gtg act tcc aat tct ctt ttg ccc cat ccc ttt ggt 2736
262 Leu Tyr Phe Gly Val Thr Ser Asn Ser Leu Leu Pro His Pro Phe Gly
263          900          905          910
265 gag cta gag gat gta cat tat tct ggt ggt caa aat ccg ata ata tac 2784
266 Glu Leu Glu Asp Val His Tyr Ser Gly Gly Gln Asn Pro Ile Ile Tyr
267          915          920          925
269 aag ata cct aag cta cta cac caa gag gtg ctc caa aat tct gaa aca 2832
270 Lys Ile Pro Lys Leu Leu His Gln Glu Val Leu Gln Asn Ser Glu Thr
271          930          935          940
273 ttt tgt tct tct gtc ggg cgt ggc atc tca aga gaa tct ttt ctg aag 2880
274 Phe Cys Ser Ser Val Gly Arg Gly Ile Ser Arg Glu Ser Phe Leu Lys
275 945          950          955          960
277 cat ttt aat ata tat tca cct gag tat att ctt aag tca ata ttc cca 2928
278 His Phe Asn Ile Tyr Ser Pro Glu Tyr Ile Leu Lys Ser Ile Phe Pro
279          965          970          975
281 tct gat agt ggg gta gat caa gtg gtt agt gga agt gga gca ttt ggc 2976
282 Ser Asp Ser Gly Val Asp Gln Val Val Ser Gly Ser Gly Ala Phe Gly
283          980          985          990
285 ttt tca cgc ttg atg gat cta tca cca tca gaa gtt gga tat ctg gct 3024
286 Phe Ser Arg Leu Met Asp Leu Ser Pro Ser Glu Val Gly Tyr Leu Ala
287          995          1000          1005
289 ctg tgt tct gtt gca gaa agg cta tta ttt tct ata ctg agg tgg gag 3072
290 Leu Cys Ser Val Ala Glu Arg Leu Leu Phe Ser Ile Leu Arg Trp Glu
291          1010          1015          1020
293 cgg caa ttt ttg gat gaa tta gtt aac tct ctt atg gag tcc aag gat 3120
294 Arg Gln Phe Leu Asp Glu Leu Val Asn Ser Leu Met Glu Ser Lys Asp
295 1025          1030          1035          1040
297 ggt gat ctt agt gac aat aac atc gag aga gtt aaa acc aaa gct gtc 3168
298 Gly Asp Leu Ser Asp Asn Asn Ile Glu Arg Val Lys Thr Lys Ala Val
299          1045          1050          1055
301 aca aga atg ttg ctg atg cca tca aaa gtt gaa acg aat ttt cag aaa 3216
302 Thr Arg Met Leu Leu Met Pro Ser Lys Val Glu Thr Asn Phe Gln Lys
303          1060          1065          1070
305 agg aga cta agc aca ggg cct acc cgt cct tca ttt gaa gcg cta gtg 3264
306 Arg Arg Leu Ser Thr Gly Pro Thr Arg Pro Ser Phe Glu Ala Leu Val
307          1075          1080          1085
309 atc tct cat cag gat agg ttt ctt tca agt atc aaa ctc ctg cat tct 3312
310 Ile Ser His Gln Asp Arg Phe Leu Ser Ser Ile Lys Leu Leu His Ser
311          1090          1095          1100
313 gca tat act tat atc cca aaa gcc aga gct cca cct gta agc att cat 3360
314 Ala Tyr Thr Tyr Ile Pro Lys Ala Arg Ala Pro Pro Val Ser Ile His
315 1105          1110          1115          1120
317 tgc tcg gac aga aat tcg gca tac aga gtt aca gaa gaa tta cat caa 3408
318 Cys Ser Asp Arg Asn Ser Ala Tyr Arg Val Thr Glu Glu Leu His Gln
319          1125          1130          1135
321 cca tgg ctt aag aga cta tta atc ggt ttt gca cga acg tca gaa gct 3456
322 Pro Trp Leu Lys Arg Leu Leu Ile Gly Phe Ala Arg Thr Ser Glu Ala
323          1140          1145          1150
325 aat gga ccc agg aag cct aac agc ttt cca cat cct tta atc caa gaa 3504

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/519,069

DATE: 11/03/2005

TIME: 08:23:19

Input Set : A:\Seqlist.txt

Output Set: N:\CRF4\11032005\J519069.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date